CDF Data Handling

Terry Watts, Rutgers University,

for DH Group.

CDF Week

May 16, 2000

CDF Data Handling

- Organize, access, archive the data.
- Components of DH (status later): Computing system, data disks, tape archive, Datafile Catalog, AC++ IO Modules, Disk Inventory Manager & Stager, LSF batch queues, datalogger.
- Ref: cdf5310, Chep2000 papers 366-368.

Organization of Data - 1

- Datasets, in hierarchy (mostly)

  - Dataset = “collection” of events of particular “content”
  - “Collection” defined by Level 3 bits for raw and pad, by group/user AC++ filter for secondary datasets, by MC job, etc.
  - “Content” is raw, pad, other.
  - Raw data forms “streams” = combined primary datasets Stream accessed as “raw dataset”.

Organization of Data - 2

Events → Runsections → Run

- Runsections defined about every 30 sec so contain 2000 events approx.
- Calculate integrated luminosity by runsection, then sum.
- Define data quality by runsection.
- Note that dynamic trigger prescale values change on runsection boundaries.

Organization of Data - 3

- For accessing data, hierarchy is:

Dataset → Fileset → File → Runsection → Events

- All events of a given runsection in same file (not in order) (1 GB).
- Fileset is group of files forming a tape “partition” (10 GB).
- Unit of data for disk management is fileset.

Datafile Catalog

- Use a database to keep “meta-data” - Oracle at Fermi.
- Main Tables:

Dataset	Runsection
Fileset	Tape
File	Tape_pool

- Other associated records contain:
- for files, list of runsections (for lum sums and qual tests).
- for runsections, dynamic trigger prescale values.
- Browser under development, see next slide.

Datafile Catalog

- 4 slides showing web browser pages for datasets, filesets, files, runsection ranges, and runsections.

Access to Data

- DHInput and DHOutput are AC++ IO modules in CVS package DHMods. Consider DHInput (later for DHOutput)

DHInput User Guide, from http://rutpc7.fnal.gov/ratnikov/Docs

An example may be found in the CDFSOFT2, directory DHMods/bin/DHInput.
A validation executable including DHInput and EventDump is made there.

The example Tcl file used for validation job looks like:

```
talk DHInput
   include dataset "MDC-1_PAD_TOP_HIGGS_SUSY"
   show include
exit
```

- Can also specify files or directories instead.

User View of Data Access

- 2 transparencies: from Chep2000 paper 368, and bubbles diagram.

Disk Inventory Manager

- Prototype running since September 1999
- Desired client-server version under construction now, basic version ready for MDC2A, 2B, and Commissioning Run.
- Expand functionality (e.g. quotas, priorities, etc.) during next few months.

LSF Batch Queues

- LSF package from Platform Computing adopted for batch system.
- Basic commands: bsub, bpeek, bkill

  bhist, bjobs, bqueues

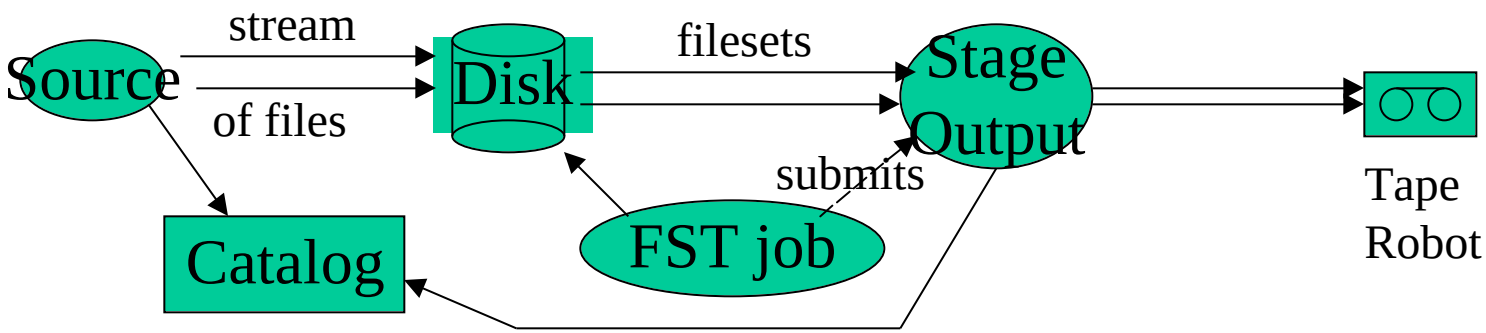
  xlsbatch
- Quotas and priorities not set up yet.
- Only basic queues: short, medium, long, very_long.

fcdfsgi2

- First big box of multi-flavor unix cluster.
- 64 processors, 4 times cdfsga, 1/3 of current goal.
- 7 TB disk here, 4TB configured for use; goal is 30 TB.
- 4 TB divided into: 2TB managed by DIM, 2TB by groups.

Archive Data

- Basic technique for adding to a dataset:



- Source (e.g. DHOutput) makes file size about 1 GB, and keeps runsections in files.
- Fileset-tape(FST) daemon forms files into filesets, and filesets into tapes.

Sources of Data

- Online: B0 CS datalogger writes a stream of files to a dual ported disk, then FST daemon takes data.
- Online: A similar low bandwidth path gets online test files to fcdfsgi2 for user access. Not archived, so no FST job.
- Farm puts out a processed stream split into the primary datasets. One FST daemon deals with all datasets.
- Group/user/MC: Proposals will exist soon for methods to create custom datasets, including MC from remote sites.

Status of Components (Short term)

• Computing	• Ready
• Disks	• Ready
• Tape archive	• Tape drives on order
• DHInput	• Ready
• DHOutput	• Updated version soon
• Catalog	• Usable, some fixes on way
• DIM/Stager	• Basic version soon
• LSF batch	• Basic setup ready
• FCC datalogger	• Updated version soon

Status of Components (Longer term)

- Computing
- Disks
- Tape archive
- DHInput
- DHOutput
- Catalog
- DIM/Stager
- LSF batch
- FCC datalogger

- Next upgrade
- Add disks
- Add drives; finalize software
- Restart, data quality, etc
- Use DIM, user catalogs
- So much work, needs a group
- Full functionality
- Replace Liz Buckley-Geer
- Needs a person

Acknowledgements

- The DH Group: Paul Hubbard, Stephan Lammel, Fedor Ratnikov, Terry Watts, Eric Wicklund, (Liz Buckley-Geer, now ex-cdf).
- Much DH support comes from the CDF Group in the Computing Division. And from computing professionals in the Computing Division (see Rick StDenis' later talk on support for the Oracle version of the Datafile Catalog).
- Discussion with other CDF people offline, online, and other, has been valuable.